

Appln. No. 09/882,472  
RCE dated Oct. 11, 2005  
Reply to Office Action of June 9, 2005  
Docket No. 6169-157

IBM Docket No. BOC9-2000-0016

### REMARKS/ARGUMENTS

These remarks are made in conjunction with the filing of a Request for Continued Examination. These remarks address the Advisory Action mailed September 7, 2005 (hereinafter Advisory Action), which followed Applicants' September 9, 2005, reply to the Office Action dated June 9, 2005 (hereinafter Office Action). This response is filed after the 3-month shortened statutory period, and as such, a retroactive extension of time is hereby requested. The Examiner is authorized to charge the appropriate extension fee to Deposit Account 50-0951.

At page 2 of the Office Action, Claims 1-22 were rejected under 35 U.S.C. § 102(e) as being anticipated by U.S. Patent No. 6,078,886 to Dragosh, *et al.* (hereinafter Dragosh). Applicants present herewith additional Claims 24, 25, and 26 to further emphasize certain aspects of Applicants' invention. The newly-presented claims do not introduce new matter and are fully supported in the specification. (See, e.g., Specification, p. 4, lines 13-25; p. 7, lines 4-22; p. 10, lines 23-26; and p. 12, lines 9-11.) Applicants respectfully maintain that the cited reference fails to disclose each feature of the claims.

### Claims 1 – 9 and 14-22

Independent Claims 1 and 14 are directed to a method and computer product, respectively, for processing speech audio in a network-connected client device. Each of the claims recites determining whether to process a speech grammar locally in a network-connected client device or, alternatively, to process the speech grammar in a remote speech server in the network. The determination is based on a characterization of the speech grammar.

In a portion cited in both the Office Action and Advisory Action, Dragosh describes a client's performing the feature extraction portion of a speech recognition process, the extracted features then being transferred to a server that continues the process including processing a speech grammar. Applicants respectfully emphasize, however, that Claims 1

Appln. No. 09/882,472  
RCE dated Oct. 11, 2005  
Reply to Office Action of June 9, 2005  
Docket No. 6169-157

IBM Docket No. BOC9-2000-0016

and 14 recite that it is the processing of the speech grammar that is being performed either in the network-connected client device or the remote server depending on the characterization of the speech grammar.

The distinction between extracting features and processing a speech grammar is forcefully brought home in the paragraph immediately following the portion of Dragosh cited in the Advisory Action and Office Action:

"While typically a grammar rule will be activated prior to the initiation of transmission of speech information [i.e., extracted features rather than digitized voice signals] from ASR [automatic speech recognition] client 130 to ASR server 100, rule activation could take place after some or all of the speech information to be recognized has been sent from ASR client 130 to ASR server 100. In such a circumstance, ASR server 100 would not begin speech recognition efforts until a grammar rule has been activated. Speech sent by ASR client 130 would not begin speech recognition until a grammar rule has been activated. Speech sent by ASR client 130 prior to activation of a grammar rule could be stored temporarily by ASR server to be processed by the recognizer or, alternatively, such speech could be ignored." (Col. 8, lines 55-65.)

This portion of Dragosh clarifies that the extracted features can be received either concurrently with or entirely before processing the speech grammar with the extracted features, but regardless, the speech grammar is processed by the server. The option of waiting for the server to receive all the extracted features is possible because feature extraction is a front-end procedure that invariably precedes processing speech according to a particular speech grammar; the speech grammar must be applied to the extracted features. Dragosh explicitly teaches that, even if front-end feature extraction is performed in a client, it remains the task of the server to process the speech grammar – a back-end portion of the

Appln. No. 09/882,472  
RCE dated Oct. 11, 2005  
Reply to Office Action of June 9, 2005  
Docket No. 6169-157

IBM Docket No. BOC9-2000-0016

overall speech recognition. Nothing in this or the other portions of Dragosh, however, remotely suggests that the speech grammar is to be processed alternatively in a client device or a remote server, let alone that the choice is based on a characterization of the speech grammar as recited in independent Claims 1 and 14.

At page 3 of the Office Action, it is stated that Dragosh teaches the "determination of locations, including locally or remotely," of a speech grammar. Applicants again emphasize that their invention is not directed to where the grammar is located. Applicants' invention is directed to determining where the grammar is to be processed – either in a network-connected client device or a remote server, the determination being based on a characterization of the speech grammar. With Dragosh, the speech grammar is either already on the server or the server obtains the speech grammar from another server or the client, but regardless of where the speech grammar is located initially, it is the server that invariably processes the speech grammar.

Dragosh discloses that a client can either send a speech grammar to a server or identify an existing grammar from among ones already stored at the server. (Col. 4, lines 45-56.) The client also can send an identifier or IP address, such as a URL-compatible address, which identifies where the server can find a grammar on a different server. (Col. 5, lines 30-50.) Although Dragosh does not explicitly distinguish grammars according to a complexity criteria, it is noted at page 3 of the Office Action that Dragosh distinguishes a common, or "canned grammar," from other grammars. Even if this is read as an implicit characterization according to complexity, however, it pertains only to where the grammar resides before being obtained and processed by the server. The client sends the grammar to the server, identifies an existing grammar on the server, or provides an address of another server where the grammar can be accessed. But in every case, Dragosh teaches that, regardless of where the grammar is to be found, it is the sever that processes the grammar. This is counter to Applicants' invention, which provides that a determination is made as whether the speech

Appln. No. 09/882,472  
RCE dated Oct. 11, 2005  
Reply to Office Action of June 9, 2005  
Docket No. 6169-157

IBM Docket No. BOC9-2000-0016

grammar is to be processed at the client device or, alternatively, on the server depending on a characterization of the grammar, as recited in independent Claims 1 and 14.

Applicants respectfully maintain that Dragosh fails to expressly or inherently teach the features of independent Claims 1 and 14, and that, therefore, the claims define over the prior art. Applicants further respectfully maintain that whereas claims 2-9 and 15-22 each depend from one of Claims 1 or 14 while recited additional features, these claims likewise define over the prior art.

#### Claims 10 – 13

Independent Claim 10 is directed to a network distributable speech grammar configured for distribution to network connected client devices. The speech grammar includes a pre-determined characterization and is configured for selectively specifying a pre-determined preference for processing the speech grammar either locally or remotely. As already noted, Dragosh fails to expressly or inherently teach a speech grammar that includes a characterization for determining whether to process the speech grammar locally or remotely. In all instances with Dragosh, the speech grammar is processed remotely.

Features may be extracted from a client, but it is the server, either while receiving the extracted features or after having received all the extracted features from the client, which processes the speech grammar. The speech grammar must be applied to the particular features, but, as already pointed out, extracted features are not synonymous with a speech grammar, nor is the extraction of features equivalent to processing a speech grammar.

It follows that Dragosh does not expressly or inherently teach every feature recited in Claim 10. Accordingly, Applicants respectfully maintain that Claim 10 defines over the prior art. Applicants also respectfully maintain that whereas claims 11-13 each depend from Claim 10 while reciting additional features, these claims similarly define over the prior art.

Appln. No. 09/882,472  
RCE dated Oct. 11, 2005  
Reply to Office Action of June 9, 2005  
Docket No. 6169-157

IBM Docket No. BOC9-2000-0016

### Claim 23

Claim 23 is directed to a method for processing speech audio in a network connected client device. The method includes selecting a speech grammar for use in a speech recognition system in the network connected client device. Additionally, the method includes characterizing the selected speech grammar based upon whether a predefined processing complexity of the speech grammar exceeds a predetermined threshold processing capability of the network connected client device. The method further includes processing the speech grammar remotely in a speech server in the network if the predefined processing complexity exceeds the predetermined threshold.

Dragosh, as previously noted, does not explicitly distinguish speech grammars on the basis of the complexity. The reference to a "common" or "canned grammar" in Dragosh does not pertain to a relative complexity of the particular grammar. Instead, as Dragosh makes explicit, a common or canned grammar is one "which ASR server 100 would already have stored." (Col. 5, lines 32-38.) It would be counterintuitive to assume that a common grammar is a complex grammar, and Dragosh resolves any doubt by stating that a common grammar is a grammar such as a simple TIME-OF-DAY or DATE grammar.

Yet even if the term common is meant to imply a lack of complexity, that would obviate any suggestion that Dragosh determines where to process a grammar based on complexity since the common grammar in Dragosh is processed by the server on which it resides, not at the client device. Were Dragosh to determine which of the two should process a particular speech grammar, it would be the client device that would process the less complex speech grammars. Dragosh, however, can not logically be inferred as implying such an approach.

More fundamentally, nothing in Dragosh remotely suggests that, even if speech grammars are characterized according to their complexity, complexity determines whether the speech grammar is processed locally on a client or remotely on a server. This follows inevitably since Dragosh only discloses processing speech grammars on a server, regardless

Appln. No. 09/882,472  
RCE dated Oct. 11, 2005  
Reply to Office Action of June 9, 2005  
Docket No. 6169-157

IBM Docket No. BOC9-2000-0016

of the location from which the grammars are obtained. This is true in Dragosh even if feature extraction is performed by a client device. Applicants respectfully submit, therefore, that Dragosh fails to teach or suggest newly-presented Claim 23 and that the claim defines over the prior art.

### Claims 24 and 25

Newly presented Claim 24 is directed to a method for processing speech audio in a network connected client device and also includes selecting a speech grammar for use in a speech recognition system in the network connected client device. The method further includes identifying a processing preference in the speech grammar based upon at least one of a predetermined complexity of the speech grammar and processing resources of the network connected client device. In this manner, a designer of the speech grammar could include in the grammar a tag, for example, that specifies that certain devices should rely on a network server for processing the speech grammar when speech processing is performed with the device. (See, e.g., Specification, p. 7, lines 11-14.) Additionally, the method includes determining whether to process the speech grammar locally in the network connected client device or remotely in a speech server in the network based on the predetermined preference.

As previously asserted Dragosh does not provide any explicit mechanism for characterizing a speech grammar, let alone characterizing one on the basis of complexity and/or relative resources of a network connected client device that might utilize the speech processing grammar. Moreover, Dragosh provides no mechanism whereby a determination is made, let alone one based on a processing preference, as to whether to process the speech grammar locally in the network client device or, alternatively, in a remote server. Applicants, therefore, respectfully submit that independent Claim 24 also defines over the prior art. Applicants further respectfully submit that Claim 25, which depends from Claim 24 and recites additional features, also defines over the prior art.

Appln. No. 09/882,472  
RCE dated Oct. 11, 2005  
Reply to Office Action of June 9, 2005  
Docket No. 6169-157

IBM Docket No. BOC9-2000-0016

### CONCLUSION

Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Date: October 11, 2005



Gregory A. Nelson, Registration No. 30,577  
Richard A. Hinson, Registration No. 47,652  
Marc A. Boillot, Registration No. 56,164  
AKERMAN SENTERFITT  
Customer No. 40987  
Post Office Box 3188  
West Palm Beach, FL 33402-3188  
Telephone: (561) 653-5000